

IB 2004/050297

PHAT030013

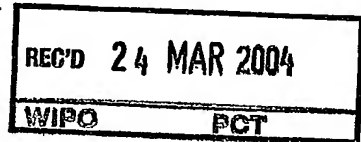


Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

BEST AVAILABLE COPY



Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03100782.6

✓

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk



Europäisches
Patentamt

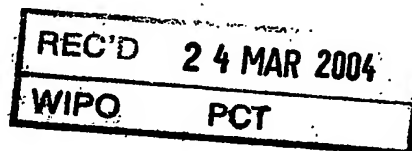
European
Patent Office

Office européen
des brevets

PCT / IB04/050294

Anmeldung Nr:
Application no.: 03100782.6 ✓
Demande no:

Anmeldetag:
Date of filing: 26.03.03 ✓
Date de dépôt:



Anmelder/Applicant(s)/Demandeur(s):

Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se référer à la description.)

Spracherkennungs- und Korrektursystem

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s)
revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

G10L/

Am Anmeldetag benannte Vertragsstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT SE SI SK TR LI

Spracherkennungs- und Korrektursystem

Die Erfindung bezieht sich auf ein Spracherkennungs- und Korrektursystem,
5 das zumindest eine Spracherkennungseinrichtung, der ein gesprochener Text zuführbar und
in einen erkannten Text transkribierbar ist, sowie eine mit der zumindest einen
Spracherkennungseinrichtung über ein Datennetz zur Übertragung des erkannten Textes
und gegebenenfalls des gesprochenen Textes verbundene Korrektureinrichtung zum
Korrigieren des von der zumindest einen Spracherkennungseinrichtung erkannten Textes
10 umfasst.

Die Erfindung bezieht sich weiters auf eine Korrektureinrichtung zum
Korrigieren eines von einer Spracherkennungseinrichtung erkannten Textes.

Die Erfindung bezieht sich weiters auf ein Alternativen-Lexikon-
Erstellungsverfahren zum Ermitteln von Datensatzeinträgen für ein Alternativen-Lexikon
15 für die Korrektur von erkanntem Text, der von einer Spracherkennungseinrichtung aus
gesprochenem Text transkribiert wurde.

Ein solches Spracherkennungs- und Korrektursystem ist aus dem Dokument
20 US 5,864,805 bekannt. Dieses Dokument offenbart ein kontinuierlich arbeitendes
Spracherkennungssystem mit der Fähigkeit, Fehler in Worten und Wortfolgen zu erkennen
und zu korrigieren. Zur Korrektur von Fehlern werden Daten im internen Speicher des
Spracherkennungssystems gespeichert, um damit im Spracherkennungssystem geführte
Wahrscheinlichkeitstabellen zu aktualisieren, die bei der Entwicklung von Alternativen-
25 Listen für die Ersetzung von falsch erkanntem Text dienen.

Bei dem bekannten Spracherkennungs- und Korrektursystem hat sich als
Nachteil erwiesen, dass es bloß als Stand-alone-Lösung benutzbar ist, das heißt, dass dieses
Spracherkennungs- und Korrektursystem auf einen Einzelplatz-Computer beschränkt ist, in
dem sämtliche Daten gespeichert sind, die das Spracherkennungs- und Korrektursystem
30 benötigt. Moderne Spracherkennungssysteme sind mittlerweile jedoch oftmals als verteilte
Systeme ausgeführt, bei denen eine Vielzahl von Computern mit darauf laufender
Spracherkennungssoftware oder Teilen davon über ein Datennetzwerk miteinander

- verbunden sind. Oftmals ist bei diesen fortgeschrittenen Systemen auch eine Verteilung der Aufgaben des Spracherkennungs- und Korrektursystems auf mehrere Computer vorgesehen. Als Beispiel dafür sei ein Spracherkennungssystem erwähnt, wie es bei der klinischen Diagnose in Krankenhäusern verwendet wird. Dabei werden von vielen Ärzten
- 5 in verschiedenen Untersuchungszimmern Diagnosen in das Spracherkennungssystem diktiert, die vom Spracherkennungssystem in einen erkannten Text umgewandelt und zusammen mit einer Audio-Aufzeichnung des gesprochenen Textes zentral gespeichert werden. Bei dem erkannten Text handelt es sich allerdings noch um eine Rohfassung, die in einem Korrekturvorgang von allfälligen Erkennungsfehlern bereinigt werden muss.
- 10 Diese Korrektur wird meist von einer Sekretärin durchgeführt, wobei es üblich ist, dass eine einzige Sekretärin die Diktate einer Vielzahl von Ärzten korrigiert. Da bei diesem Spracherkennungssystem sowohl die Ärzte in den einzelnen Untersuchungszimmern als auch die Sekretärin in einem Büro voneinander entfernt und meist auch zu unterschiedlichen Zeiten arbeiten, ist die in dem Dokument US 5,864,805 vorgeschlagene
- 15 Lösung für ein verteiltes Spracherkennungssystem nicht anwendbar. Andererseits ist es auch nicht praktikabel, die beim Transkriptionsvorgang des Spracherkennungssystems anfallende Information, die zur Erstellung von Alternativen-Listen für die Korrektur benutzt werden könnte, über ein Datennetzwerk an jenen Computer zu übertragen, auf dem die Korrektur des erkannten Textes durchgeführt werden soll, da die anfallenden
- 20 Datenmengen viel zu hoch wären. So würden die in dem Dokument US 5,864,805 erwähnten Wahrscheinlichkeitstabellen viel zu schnell anwachsen, um in ständiger Aktualisierung über ein Datennetzwerk an eine Korrekturereinrichtung übertragen zu werden, insbesondere wenn es sich um ein Datennetzwerk mit geringer Bandbreite handelt. Es ist in der Praxis auch nicht durchführbar, die beim Transkriptionsvorgang des
- 25 Spracherkennungssystems anfallende Information direkt an eine Korrekturereinrichtung zu übertragen und die Information dort zu analysieren, da auch in diesem Fall die benötigten Netzwerkbandbreiten, speziell für Netzwerke mit geringer Bandbreite viel zu hoch wären. Es ist nämlich dabei zu bedenken, dass moderne Spracherkennungssysteme typischerweise 5.000 bis 8.000 Wahrscheinlichkeitshypothesen, wie ein gesprochener Text in erkannten
- 30 Text umgewandelt werden könnte, parallel bearbeiten. Information aus diesen Wahrscheinlichkeitshypothesen wäre aber für die Korrekturereinrichtung erforderlich. Wenn nun beispielsweise ein Erkennungsergebnis, d.h. die beste Hypothese aus 1000 Wörtern

besteht und jedes Wort zehnmal im zugrundeliegenden Wortgraph vorkommt, müsste man im Extremfall Varianten übertragen, die aus 1000 hoch 10 Wörtern bestehen und sich nur durch unterschiedliche zeitliche Aufteilung unterscheiden.

Andererseits ist, obwohl die Entwickler von Spracherkennungssystemen intensiv an Verbesserungen ihrer Systeme arbeiten, auf absehbare Zeit noch keine 100%-ige Erkennungsrate zu erwarten, so dass weiterhin Korrekturen am erkannten Text erforderlich sein werden. Es besteht daher ein Bedürfnis, diese Korrekturarbeit zu erleichtern, indem der korrigierenden Person während der Korrektur Alternativen zu falsch erkannten Wörtern angeboten werden, so dass sie eine der angebotenen Alternativen rasch auswählen kann.

Die Erfindung hat sich zur Aufgabe gestellt, ein Spracherkennungs- und Korrektursystem gemäß der in dem ersten Absatz angegebenen Gattung, eine Korrekturereinrichtung gemäß der in dem zweiten Absatz angegebenen Gattung und ein Alternativen-Lexikon-Erstellungsverfahren gemäß der in dem dritten Absatz angegebenen Gattung zu schaffen, wobei die vorstehend erläuterten Nachteile vermieden sind. Zur Lösung vorstehend angegebener Aufgabe ist bei einem solchen Spracherkennungs- und Korrektursystem vorgesehen, dass die Korrekturereinrichtung ein Alternativen-Lexikon aufweist, das Wortteile, Wörter bzw. Wortfolgen enthält, die von der Korrekturereinrichtung als Alternativen zu einzelnen Wortteilen, Wörtern bzw. Wortfolgen des erkannten Textes anzeigbar sind.

Zur Lösung vorstehend angegebener Aufgabe ist bei einer solchen Korrekturereinrichtung vorgesehen, dass in der Korrekturereinrichtung ein Alternativen-Lexikon gespeichert ist, das Wortteile, Wörter bzw. Wortfolgen enthält, die von der Korrekturereinrichtung als Alternativen zu einzelnen Wortteilen, Wörtern bzw. Wortfolgen des erkannten Textes anzeigbar sind.

Der Begriff „Alternativen-Lexikon“ ist dabei so zu verstehen, dass es auf Informationen basiert, die vom Transkriptionsvorgang von Spracherkennungseinrichtungen unabhängig sind. Insbesondere basiert das Alternativen-Lexikon nicht auf alternativen Erkennungshypothesen, die von Spracherkennungseinrichtungen während des Transkriptionsvorgangs auch aufgestellt und bezüglich der Wahrscheinlichkeit ihres

Zutreffens schlechter bewertet wurden als die sich im erkannten Text widerspiegelnde Erkennungshypothese.

Zur Lösung vorstehend angegebener Aufgabe ist bei einem solchen Alternativen-Lexikon-Erstellungsverfahren vorgesehen, dass von der

- 5 Spracherkennungseinrichtung unabhängige Wissensquellen, insbesondere für das Anwendungsgebiet spezifische Textdateien, wie medizinische oder juristische Texte, oder aus einer Vielzahl von korrigierten Texten und zugehörigen von Spracherkennungseinrichtungen erzeugten erkannten Texten (ET) erstellte Verwechslungsstatistiken, auf miteinander verwechselbare Textelemente, wie Wortteile,
- 10 Wörter bzw. Wortfolgen untersucht werden, und solche miteinander verwechselbaren Textelemente als Alternativen in einem Datensatzeintrag zusammengestellt werden.

- Durch die erfindungsgemäßen Merkmale ist erreicht, dass die Korrektur von durch ein Spracherkennungssystem erkannten Texten wesentlich einfacher und schneller als bisher durchgeführt werden kann, wobei die Erfindung besonders vorteilhaft bei
- 15 Spracherkennungssystemen eingesetzt werden kann, bei denen Erkennung und Korrektur nicht auf demselben Computer durchgeführt wird. Die durch die Erfindung erzielte Angabe von Korrekturalternativen ist darüber hinaus überaus effizient, flexibel und robust, d.h. unabhängig von spezifischen Erkennungsfehlern. Neben dem Entfall eines umfangreichen Datentransfers zwischen Spracherkennungssystem und Korrekturereinrichtung bei der
- 20 Korrektur bietet die Erfindung darüber hinaus den weiteren großen Vorteil, dass die Alternativvorschläge aus dem Alternativen-Lexikon unabhängig von der jeweiligen Erkennungsleistung der Spracherkennungseinrichtung sind. Dem gegenüber bestand bei bisherigen Systemen der Nachteil, dass im Falle von geringer Erkennungsrate der Spracherkennungseinrichtung vielfach auch bei der Korrektur keine brauchbaren
- 25 Alternativen angeboten wurden, da diese ebenfalls falsch waren.

- Gemäß den Maßnahmen des Anspruchs 2 ist der Vorteil erhalten, dass die Korrekturereinrichtung unabhängig von der während des Transkriptionsvorgangs in dem Spracherkennungssystem anfallenden Information betreibbar ist, so dass abgesehen von der Übertragung des erkannten Textes und gegebenenfalls des zugrundeliegenden
- 30 gesprochenen Textes keine Datenkommunikation zwischen dem Spracherkennungssystem und der Korrekturereinrichtung erforderlich ist. Durch die große Flexibilität der erfindungsgemäßen Lösung ist auch eine einfache Adaption an neue Kontexte oder

Sprechstile erzielbar. Die Korrekturereinrichtung kann sich in einer bevorzugten Ausgestaltung auf Analysemittel zum Analysieren von selektierten Textstellen des erkannten Textes stützen, die vorzugsweise mittels Zeichenkettenvergleich oder höheren syntaktischen Analyseverfahren Alternativen zu den selektierten Textstellen aus dem Alternativen-Lexikon ermitteln. Syntaktische Analyseverfahren umfassen beispielsweise die Erfassung von syntaktischen Konstituenten, wie Nomen-Verb-Paaren, Nominalphrasen etc.

Gemäß den Maßnahmen des Anspruchs 4 ist der Vorteil erhalten, dass der Benutzer, beispielsweise über definierte Hot-Keys an der Tastatur der Korrekturereinrichtung, sich Alternativen zu den gerade bearbeiteten Stellen des erkannten Textes anzeigen lassen kann.

Gemäß den Maßnahmen des Anspruchs 5 wiederum ist der Vorteil erhalten, dass die Korrekturereinrichtung über im Hintergrund laufende Analysemittel ständig von sich aus Alternativen zu selektierten Textstellen anbieten kann.

Gemäß den Maßnahmen des Anspruchs 7 ist der Vorteil erhalten, dass das Alternativen-Lexikon sowohl offline wie auch online unabhängig von einem Spracherkennungssystem erstellt und aktualisiert werden kann, da die Wissensquellen, auf die zurückgegriffen wird, unabhängig von jener Information sind, die beim Transkriptionsvorgang eines Spracherkennungssystems vorübergehend oder ständig zur Verfügung steht.

Gemäß den Maßnahmen des Anspruchs 8 ist der Vorteil erhalten, dass das Wissen zur Ermittlung von Datensatzeinträgen aus der Erfahrung bzw. der Beobachtung stammt, welche Verwechslungen von einem Spracherkennungssystem besonders häufig gemacht werden. Beispielsweise werden naturgemäß homophone Wörter, das heißt gleich klingende, aber unterschiedlich geschriebene Wörter vom Spracherkennungssystem besonders häufig verwechselt. Durch das Heranziehen der Korrekturinformation zur Erstellung eines Alternativen-Lexikons kann gleichsam die Leistung eines Spracherkennungssystems nachträglich verbessert werden, ohne dass es erforderlich ist, das Spracherkennungssystem auf die gemachten Fehler erneut zu trainieren. Anders gesagt lernt das Korrektursystem aus den Fehlern des Spracherkennungssystems.

Um die Robustheit des Verfahrens zur Erstellung von Datensatzeinträgen in einem Alternativen-Lexikon zu erhöhen, können vorteilhaft statistische Verfahren

herangezogen werden, wie in den Ansprüchen 8 bis 10 dargelegt. Durch diese statistischen Verfahren wird einerseits erreicht, dass die Alternativen-Liste für ein zu ersetzendes Wortelement nicht zu viele Einträge bekommt und damit für den Benutzer unhandlich wird, indem nur solche Alternativen aufgenommen werden, die hinreichend oft in der

5 Korrektur vorkommen. Andererseits wird durch die Einführung eines oberen Grenzwerts für die Häufigkeit einer Ersetzung während der Korrektur sichergestellt, dass systematische Ersetzungen, die (fast) immer durch dasselbe Wortelement korrigiert werden, wie z.B. die Anweisung „Briefende“ in einem Diktat durch „mit freundlichen Grüßen, Dr. Meyer“ nicht als einzige Alternative angeboten werden. Ein solcher Fall sollte durch andere

10 Mechanismen geregelt werden.

Gemäß den Maßnahmen des Anspruchs 11 ist der Vorteil erhalten, dass erkannt wird, ob es sich um eine „ernst zu nehmende“ Ersetzung handelt, z.B. die Ersetzungen von ‚mein – dein‘, ‚dem – den‘ etc. Die erforderliche phonetische Ähnlichkeit kann entweder über den gesprochenen Text, der in diesem Fall an die Korrektureinrichtung übermittelt

15 wird, oder aus der Lautschrift der betrachteten Wörter, die der Korrektureinrichtung bekannt ist, ermittelt werden.

Gemäß den Maßnahmen des Anspruchs 12 ist der Vorteil erhalten, dass in die Alternativen-Liste nur solche Wortelemente aufgenommen werden, die zeitmäßig etwa an derselben Stelle im gesprochenen Text zu liegen kommen. Wenn beispielsweise

20 systematisch vom Benutzer bei der Korrektur irgendwelche Wörter oder Textabschnitte eingefügt werden, die nichts mit dem gesprochenen Text zu tun haben oder systematisch Wörter ausgelassen werden, die dann nicht im korrigierten Text aufscheinen, ist es nicht zweckmäßig, solche Korrekturen über Alternativen-Listen abzuhandeln.

Datensatzeinträge in einem Alternativen-Lexikon können unterschiedliche

25 Detaillierungsgrade haben. So können unterschiedliche Alternativen-Listen je nach verwendeter Sprache im gesprochenen Text erstellt werden. Weiters können die Datensatzeinträge des Alternativen-Lexikons nach Fachgebieten bzw. Anwendungsgebieten unterteilt werden oder nach Autoren des zugrundeliegenden gesprochenen oder korrigierten Textes unterteilt werden. Es sind auch Kombinationen der

30 oben erwähnten Detaillierungsgrade möglich.

Gemäß den Maßnahmen des Anspruchs 17 ist der Vorteil erhalten, dass das Alternativen-Lexikon während der Korrektur von erkannten Texten laufend verbessert wird

und sich somit selbst trainiert.

- Die Erfindung wird im Folgenden anhand von einem in der Figur dargestellten
- 5 Ausführungsbeispiel beschrieben, auf das die Erfindung aber nicht beschränkt ist.

Die Figur 1 zeigt ein Spracherkennungssystem mit einem damit über ein Datennetzwerk verbundenen Korrektursystem.

- 10 Die Figur 1 zeigt eine Spracherkennungseinrichtung 1 zum Transkribieren eines gesprochenen Textes GT. Die Spracherkennungseinrichtung 1 kann durch einen Computer gebildet sein, der eine Spracherkennungssoftware abarbeitet. Die Spracherkennungseinrichtung 1 umfasst Spracherkennungsmittel 7, Parameterspeichermittel 9, Befehlsspeichermittel 10 und eine Adaptionstufe 11. Über ein
- 15 Mikrophon 5 ist ein gesprochenen Text GT repräsentierendes Audiosignal A an einen A/D-Wandler 6 abgebar, der das Audiosignal A in digitale Audiodaten AD wandelt, die den Spracherkennungsmitteln 7 zuführbar sind. Die digitalen Audiodaten AD werden von den Spracherkennungsmitteln 7 in erkannten Text ET umgewandelt, der über ein Datennetzwerk 2 in Speichermitteln 8 gespeichert wird. Dazu werden in den
- 20 Parameterspeichermitteln 9 gespeicherte Parameterinformationen PI berücksichtigt, die Wortschatzinformationen, Sprachmodellinformationen und Akustikinformationen enthalten.

- Die Wortschatzinformationen beinhalten sämtliche mit den Spracherkennungsmitteln 7 erkennbaren Wörter samt zugehörigen Phonemfolgen. Die
- 25 Wortschatzinformationen werden durch die Analyse einer Vielzahl von für die vorgesehene Anwendung relevanten Texten gewonnen. Beispielsweise werden für ein Spracherkennungssystem, das auf dem Gebiet der Radiologie eingesetzt wird, Befunde mit einer Gesamtanzahl von 50 bis 100 Millionen enthaltenen Wörtern analysiert. Die Sprachmodellinformationen enthalten statistische Informationen zu in der Sprache des
- 30 gesprochenen Textes GT üblichen Abfolgen von Wörtern, insbesondere die Wahrscheinlichkeiten für das Auftreten von Wörtern und ihren Zusammenhang mit vor- und nachstehenden Wörtern. Die Akustikinformationen enthalten Informationen über die

Eigenarten der Aussprache eines Benutzers der Spracherkennungseinrichtung 1 und über akustische Eigenschaften des Mikrophons 5 und des A/D-Wandlers 6.

In dem Dokument US 5,031,113, dessen Offenbarung durch diesen Hinweis als in die Offenbarung dieses Dokuments aufgenommen gilt, ist die Abarbeitung eines

5 Spracherkennungsverfahrens unter Berücksichtigung solcher Parameterinformationen PI offenbart, weshalb hierauf nicht näher eingegangen wird. Als Ergebnis des Spracherkennungsverfahrens sind von den Spracherkennungsmitteln 7 den erkannten Text ET enthaltende Textdaten in den Speichermitteln 8 speicherbar. Weiters ist in den Speichermitteln 8 der gesprochene Text GT in digitalisierter Form speicherbar. Zusätzlich

10 wird von dem Spracherkennungssystem 1 Information über die verwendete Sprache 14, die Anwendung 15 und den Autor 16 zusammen mit dem erkannten Text ET über das Datennetzwerk 2 übertragen und in den Speichermitteln 8 abgespeichert.

In der Befehlsspeicherstufe 10 sind Abfolgen von Worten gespeichert, die von den Spracherkennungsmitteln 7 als Kommando erkannt werden. Solche Kommandos

15 umfassen beispielsweise die Abfolge der Worte „next word bold“, um das nächste Wort in dem erkannten Text ET fett darzustellen.

Eine Korrektureinrichtung 3 hat Zugriff auf den in den Speichermitteln 8 abgespeicherten erkannten Text ET, um ihn zusammen mit der akustischen Information über den zugrundeliegenden gesprochenen Text GT und die Informationen über die

20 Sprache 14, die Anwendung (das Fachgebiet) 15 und den Autor 16 einzulesen, so dass der erkannte Text ET mit einem Textverarbeitungssystem korrigiert werden kann. Insbesondere sind auf den erkannten Text ET alle Funktionen anwendbar, über die fortgeschrittene Spracherkennungssoftware verfügt, wie unten erläutert wird. Die Korrektureinrichtung 3 verfügt über Wiedergabe- und Korrekturmittel 18, an die eine

25 Tastatur 19, ein Monitor 20 und ein Lautsprecher 21 angeschlossen sind. Die Wiedergabe- und Korrekturmittel 18 sind zum visuellen Darstellen des erkannten Textes ET am Monitor 20 sowie zum akustischen Wiedergeben des gesprochenen Textes GT am Lautsprecher 21 und zum synchronen visuellen Markieren der gerade akustisch wiedergegebenen Stellen des gesprochenen Textes im erkannten Text ET ausgebildet, wenn sich die Wiedergabe-

30 und Korrekturmittel 18 in aktivierter synchroner Wiedergabebetriebsart befinden. In dieser Wiedergabebetriebsart ist gleichzeitig das Korrigieren des erkannten Textes mittels Tastatureingabe sowie gegebenenfalls mittels Sprachbefehlen über ein nicht dargestelltes

Mikrophon möglich. Der korrigierte Text KT ist in Speichermitteln 17 speicherbar.

Die Wiedergabe- und Korrekturmittel 18 umfassen Analysemittel 24 zum Analysieren von gerade selektierten Textstellen des erkannten Textes ET, um dem Benutzer der Korrektureinrichtung Alternativen zu den selektierten Textstellen aus einem Alternativen-Lexikon 23 vorzuschlagen, das in der Korrektureinrichtung 3 gespeichert ist. Die Alternativen werden dem Benutzer in Form einer Alternativen-Liste 22 am Monitor präsentiert, in der er mittels Cursor-Tasten der Tastatur 19 oder einer nicht dargestellten Maus oder dergleichen navigieren kann, um seine Auswahl zur Ersetzung zu treffen oder aber händisch eine Korrektur vorzunehmen. Die Analysemittel 24 arbeiten entweder kontinuierlich im Hintergrund oder sind vom Benutzer der Korrektureinrichtung 3 durch das Drücken einer Taste oder Tastenkombination („Hotkey“) aktivierbar. Die Analysemittel analysieren die selektierten Textstellen vorzugsweise entweder durch Zeichenkettenvergleich oder ein syntaktisches Analyseverfahren. Der Zeichenkettenvergleich kann auf Einzelwörtern oder Bestandteilen von Einzelwörtern, oder auf (zu detektierenden) Phrasen beruhen. Der Vergleich kann weiters auf Ausdrücken aus mehreren syntaktischen Konstituenten, wie z.B. Nomen-Verb-Paare, Nominalphrasen etc. beruhen. All diese Ausdrücke werden in der vorliegenden Patentanmeldung auch allgemein mit dem Begriff Textelement bezeichnet. Die von den Analysemitteln 24 vorgeschlagenen Alternativen-Listen können wiederum Einzelworte oder Teile davon, oder ganze Phrasen umfassen. Als Beispiel für die Ersetzung von Einzelworten sei der folgende erkannte Text genannt: „The epigastric vessels were seen *interiorly*, and he had history of *edema*.“ Die kursiv geschriebenen Wörter wurden fehlerkannt. Wenn sich nun der Cursor bei der Korrektur dieses erkannten Textes über dem Wort „interiorly“ befindet, wird von den Analysemitteln eine Alternativen-Liste zur Korrektur angeboten, die einen oder mehrere Einträge enthält, darunter auch das Wort „anteriorly“, das in diesem Fall das korrekte Wort wäre. Der Benutzer kann durch einfache Auswahl des Wortes „anteriorly“ die schnelle Ersetzung von „interiorly“ durch dieses Wort veranlassen. Dasselbe gilt für das Wort „edema“, das durch das in einer weiteren Alternativen-Liste angebotene Wort „anemia“ zu ersetzen ist. Der Benutzer kann somit mit wenigen Tastendrücken den fehlerkannten Satz zu: „The epigastric vessels were seen anteriorly, and he had history of anemia.“ richtig stellen. In einem Beispiel für die Ersetzung von Phrasen kann als Alternative für die erkannte Phrase „*rhythm without lists*“ die korrekte Phrase „rhythm without lifts“

angeboten werden. Es ist zu beachten, dass in diesem Fall, obwohl im erkannten Text bei der Ersetzung scheinbar nur ein Buchstabe geändert wird, die gesamte zitierte Phrase untersucht, als Alternative angeboten und bei entsprechender Auswahl durch den Benutzer der Korrektureinrichtung ersetzt wird. Ein weiteres Beispiel betrifft Alternativen mit

5 mehreren Konstituenten. Diese Konstituenten können Fachausdrücke, Nomen-Verb-Paare, etc. sein. Die Analysemittel 24 können dabei einen Algorithmus anwenden, bei dem zunächst ein Tagging des erkannten Textes und die Berechnung von Konfidenzmaßen für die einzelnen Wörter (Elemente) durchgeführt wird. Für Nomen mit niedriger Konfidenz wird anschließend ein Nomen-Verb-Paar oder eine Nominalphrase ermittelt. Daran

10 anschließend wird für das komplementäre Element (Verb oder Nomen) die Identität der zugehörigen Alternativenliste durch einen Zeichenkettenvergleich ermittelt, worauf die Anzeige der anderen Elemente in Form einer Alternativen-Liste erfolgt. Durch dieses Verfahren kann beispielsweise im fehlerkannten Satz: „The extraneous tendinous *materials* were all debrided.“ die Richtigstellung des Wortes „materials“ durch „trails“ in einer

15 Alternativen-Liste angeboten werden, indem die Analysemittel 24 das niedrige Konfidenzmaß des Wortes „materials“ feststellen, die Identifikation des Nomen-Verb-Paars „material debrided“ durchführen und über das Verb „debrided“ eine Bestimmung der relevanten Alternativen-Liste vornehmen, in der der Eintrag „trails debrided“ aufscheint. Falls dieser Eintrag vom Benutzer ausgewählt wird, so werden das Nomen und das Verb

20 ersetzt, auch wenn für den Benutzer im Text nur die Ersetzung von „materials“ durch „trails“ sichtbar wird. Als weiteres Beispiel für die Ermittlung von Alternativen mit mehreren Konstituenten sei der fehlerkannte Satz: „Discharge medications two CCU“ angeführt, der eigentlich als „Disposition to CCU“ erkannt werden hätte sollen. Die Analysemittel 24 stellen die niedrige Konfidenz des Wortes „medications“ fest und

25 identifizieren die Nominalphrase „Discharge medications“. Die Bestimmung eines relevanten Eintrags in der Alternativen-Liste erfolgt über den Begriff „CCU“ und lautet „Disposition to CCU“. Dieser Eintrag kann vom Benutzer ausgewählt werden und ersetzt den gesamten, oben angeführten, fehlerkannten Satz.

Die Analysemittel 24 ermitteln selektierte Stellen des erkannten Textes ET

30 beispielsweise aus der Cursorposition eines Textverarbeitungsprogramms, das zur Korrektur des erkannten Textes benutzt wird, oder aus der Zeitposition der gerade gesprochenen Textstelle und deren Verknüpfung mit dem erkannten Text. Es ist für den

Benutzer der Korrektureinrichtung 3 somit eine effektive und schnelle Korrektur des erkannten Textes durch Auswahl von Alternativformulierungen möglich.

- Die Korrektureinrichtung 3 umfasst auch Auswertemittel 4 zur Erstellung des Alternativen-Lexikons 23 bzw. einzelner Einträge davon. Es ist anzumerken, dass
- 5 Auswertemittel auch unabhängig von der Korrektureinrichtung 3 vorgesehen sein können, um aus verschiedenen, von dem Spracherkennungssystem 1 unabhängigen Wissensquellen ein Basis-Alternativen-Lexikon zu erstellen, das dann in der Korrektureinrichtung 3 zur Verwendung gespeichert werden kann. Im dargestellten Ausführungsbeispiel greifen die Auswertemittel 4 über die Wiedergabe- und Korrekturmittel 18 auf die in den
- 10 Speichermitteln 8 gespeicherten gesprochenen und erkannten Texte GT, ET sowie auf die Information über die Sprache 14, die Anwendung 15 und den Autor 16 zu, wobei in einer alternativen Ausgestaltungsform auch ein direkter Zugriff der Auswertemittel 4 auf die Speichermittel 8 realisiert sein kann. Die Auswertemittel 4 lesen weiters den korrigierten Text KT aus den Speichermitteln 17 ein, um ihn mit dem erkannten Text ET zu
- 15 vergleichen und somit die an dem erkannten Text ET durchgeführten Textelement-Ersetzungen zu ermitteln. Diese Textelement-Ersetzungen werden statistisch analysiert und als Alternativen in Datensatzeinträge des Alternativen-Lexikons 23 aufgenommen, wenn sie bestimmte, weiter unten näher erläuterte Bedingungen erfüllen. So wird in einer bevorzugten Ausgestaltungen der erkannte Text ET mit dem korrigierten Text KT
- 20 verglichen und jene Ersetzungen ermittelt, die über den gesamten Text die geringste Gesamtabweichung, d.h. die minimale Fehleranzahl ergeben. Mit dieser Information wird die Alternativen-Liste erstellt. Zur Verbesserung der Robustheit des Systems, d.h. um die Alternativen-Liste nicht mit zu vielen Einträgen zu füllen, ist es sinnvoll, eine Statistik zu erstellen, wie häufig ein Wortelement durch ein anderes ersetzt wird. Erst wenn ein
- 25 vorgegebener unter Grenzwert der relativen oder absoluten Häufigkeit überschritten wird, wird das ersetzende Wortelement in die Alternativen-Liste aufgenommen. Es kann auch sinnvoll sein, einen oberen Grenzwert der Häufigkeit für die Ersetzung eines Wortelements durch ein anderes einzuführen, der unterschritten werden muss, um das Wortelement in eine Alternativen-Liste aufzunehmen. Wird der obere Grenzwert überschritten, so deutet
- 30 dies entweder auf einen systematischen Fehler des Spracherkennungssystems hin, der nicht über Alternativen-Listen bereinigt werden sollte, oder auf die Ersetzung von Textteilen, die nicht aufgrund von Fehlerkennungen durchgeführt wird.

Eine weitere Maßnahme zur Verbesserung der Robustheit der Erstellung eines Alternativen-Lexikons betrifft das Analysieren der phonetischen Ähnlichkeit der ersetzten und ersetzenden Begriffe. Man kann damit sicherstellen, dass diese Begriffspaare eine ausreichende phonetische Ähnlichkeit aufweisen, z.B. mein – dein, dem – den, um als Fehlerkennung des Spracherkennungssystems angesehen zu werden, deren Aufnahme in eine Alternativen-Liste Sinn macht.

Wieder eine andere Maßnahme zur Verbesserung der Robustheit der Erstellung eines Alternativen-Lexikons betrifft das Analysieren der Zeitposition, in der die korrigierten Textelemente liegen. In eine Alternativen-Liste werden demnach nur jene Textelemente aufgenommen, die zeitmäßig etwa an der entsprechenden Stelle im gesprochenen Text zu liegen kommen. Damit kann man verhindern, dass beispielsweise Wörter, die vom Benutzer der Korrekturereinrichtung aus Formatierungs- oder inhaltlichen Gründen in den erkannten Text eingefügt oder daraus gelöscht werden, aber nichts mit der Korrektur von Fehlerkennungen zu tun haben, in Alternativen-Listen Eingang finden.

Weiters können die Datensätze in dem Alternativen-Lexikon zusätzlich nach verwendeter Sprache, Anwendung (Fachgebiet) und Autor bzw. einer Kombination daraus unterteilt werden. Sobald ein erkannter Text ET zur Korrektur gelangt, von dem zusätzlich Information über die Sprache 14, Anwendung (Fachgebiet) 15 und Autor 16 bekannt ist, wird die entsprechende Alternativen-Liste aus dem Alternativen-Lexikon 23 geladen und steht zum schnellen Abrufen von Alternativen zur Verfügung.

Die Auswertemittel 4 arbeiten in der dargestellten Ausführungsform ständig im Hintergrund, so dass das Alternativen-Lexikon 23 online verbessert und somit gleichsam trainiert wird.

Neben der Auswertung des korrigierten Textes KT zur Erstellung des Alternativen-Lexikons 23 können die Auswertemittel 4 zusätzlich oder alternativ dazu auch auf andere von dem Spracherkennungssystem 1 unabhängige Wissensquellen zurückgreifen, insbesondere auf Textdateien 12, z.B. klinische Befunde, gegebenenfalls auch Verwechslungsstatistiken, die zur Erstellung von Datensatzeinträgen im Alternativen-Lexikon 23 analysiert werden. Diese Dateien können einerseits auf einer Festplatte eines Computers gespeichert sein, auf dem die Analysemittel 4 abgearbeitet werden, andererseits kann auf solche Dateien auch über ein Datennetz zugegriffen werden. Vorteilhaft kann auch das Internet nach für die Analyse geeigneten Internet-Dateien 13 durchsucht werden,

wobei dieser Vorgang – wie überhaupt das gesamte Verfahren zur Ermittlung von Datensatzeinträgen im Alternativen-Lexikon – sich sehr gut für die Automatisierung eignet.

Patentansprüche:

1. Spracherkennungs- und Korrektursystem, das zumindest eine Spracherkennungseinrichtung (1), der ein gesprochener Text (GT) zuführbar und in einen erkannten Text (ET) transkribierbar ist, sowie eine mit der zumindest einen
- 5 Spracherkennungseinrichtung (1) über ein Datennetz (2) zur Übertragung des erkannten Textes (ET) und gegebenenfalls des gesprochenen Textes (GT) verbundene Korrekturereinrichtung (3) zum Korrigieren des von der zumindest einen Spracherkennungseinrichtung (1) erkannten Textes (ET) umfasst, dadurch gekennzeichnet,
- 10 dass die Korrekturereinrichtung (3) ein Alternativen-Lexikon (23) aufweist, das Wortteile, Wörter bzw. Wortfolgen enthält, die von der Korrekturereinrichtung (3) als Alternativen zu einzelnen Wortteilen, Wörtern bzw. Wortfolgen des erkannten Textes anzeigbar (22) sind.
2. Korrekturereinrichtung zum Korrigieren eines von einer Spracherkennungseinrichtung erkannten Textes,
- 15 dadurch gekennzeichnet, dass in der Korrekturereinrichtung (3) ein Alternativen-Lexikon (23) gespeichert ist, das Wortteile, Wörter bzw. Wortfolgen enthält, die von der Korrekturereinrichtung (3) als Alternativen zu einzelnen Wortteilen, Wörtern bzw. Wortfolgen des erkannten Textes anzeigbar (22) sind.
- 20 3. Korrekturereinrichtung gemäß Anspruch 2, gekennzeichnet durch Analysemittel (24) zum Analysieren von selektierten Textstellen des erkannten Textes, vorzugsweise mittels Zeichenkettenvergleich oder syntaktische Analyse, und zum Ermitteln von Alternativen zu den selektierten Textstellen aus dem Alternativen-Lexikon (23).
- 25 4. Korrekturereinrichtung gemäß Anspruch 3, dadurch gekennzeichnet, dass die Analysemittel (24) durch einen Benutzer der Korrekturereinrichtung aktivierbar sind.
5. Korrekturereinrichtung gemäß Anspruch 3, dadurch gekennzeichnet, dass die Analysemittel (24) selektierte Textstellen aus der Cursor-Position oder der Markierungsinformation eines Textverarbeitungsprogramms ermitteln.
- 30 6. Korrekturereinrichtung gemäß Anspruch 3, dadurch gekennzeichnet, dass die Analysemittel (24) selektierte Textstellen aus der Zeitposition des gesprochenen Textes und deren Verknüpfung mit dem erkannten Text ermitteln.

7. Alternativen-Lexikon-Erstellungsverfahren zum Ermitteln von Datensatzeinträgen für ein Alternativen-Lexikon (23) für die Korrektur von erkanntem Text (ET), der von einer Spracherkennungseinrichtung (1) aus gesprochenem Text (GT) transkribiert wurde, dadurch gekennzeichnet,

- 5 dass von der Spracherkennungseinrichtung (1) unabhängige Wissensquellen (12, 13), insbesondere für das Anwendungsgebiet spezifische Textdateien, wie medizinische oder juristische Texte, oder aus einer Vielzahl von korrigierten Texten (KT) und zugehörigen von Spracherkennungseinrichtungen erzeugten erkannten Texten (ET) erstellte Verwechslungsstatistiken, auf miteinander verwechselbare Textelemente, wie
- 10 Wortteile, Wörter bzw. Wortfolgen untersucht werden, und solche miteinander verwechselbaren Textelemente als Alternativen in einem Datensatzeintrag des Alternativen-Lexikons (23) zusammengestellt werden.

8. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 7, dadurch gekennzeichnet, dass die in einem korrigierten Text (KT) gegenüber dem
- 15 zugrundeliegenden, von einer Spracherkennungseinrichtung transkribierten erkannten Text (ET) vorgenommenen Textelement-Ersetzungen ermittelt und als Alternativen in Datensatzeinträge des Alternativen-Lexikons aufgenommen werden.

9. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 8, dadurch gekennzeichnet, dass die Häufigkeit einer jeweiligen Textelement-Ersetzung statistisch
- 20 ausgewertet und die Aufnahme als Alternative in einem Datensatzeintrag des Alternativen-Lexikons nur dann vorgenommen wird, wenn ein vorgegebener unterer Grenzwert der Häufigkeit, ausgedrückt durch die absolute Anzahl an Ersetzungen, oder das Verhältnis an Ersetzungen zur Gesamtzahl von untersuchten Wörtern oder zum Gesamtaufreten eines bestimmten Wortes, überschritten wird.

- 25 10. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 9, dadurch gekennzeichnet, dass die Häufigkeit einer jeweiligen Textelement-Ersetzung statistisch ausgewertet und die Aufnahme als Alternative in einem Datensatzeintrag des Alternativen-Lexikons nur dann vorgenommen wird, wenn ein vorgegebener oberer Grenzwert der Häufigkeit, ausgedrückt durch die absolute Anzahl an Ersetzungen oder das Verhältnis an
- 30 Ersetzungen zur Gesamtzahl von untersuchten Wörtern, unterschritten wird.

11. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 8, dadurch gekennzeichnet, dass eine Analyse der akustischen Ähnlichkeit der Textelement-

Ersetzungen vorgenommen wird und die Aufnahme als Alternative in einem Datensatzeintrag des Alternativen-Lexikons nur dann vorgenommen wird, wenn ein vorgegebenes Maß an phonetischer Ähnlichkeit gegeben ist.

5 12. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 8, dadurch gekennzeichnet, dass eine Analyse der Zeitpositionen der Textelement- Ersetzungen vorgenommen wird und die Aufnahme als Alternative in einem Datensatzeintrag des Alternativen-Lexikons nur dann vorgenommen wird, wenn für das ersetzte Textelement im zugrunde liegenden gesprochenen Text (GT) in ausreichender zeitlicher Nähe ein entsprechendes Textelement vorhanden ist.

10 13. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 7, dadurch gekennzeichnet, dass die Datensatzeinträge des Alternativen-Lexikons nach Sprachen (14) unterteilt werden.

15 14. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 7, dadurch gekennzeichnet, dass die Datensatzeinträge des Alternativen-Lexikons nach Fachgebieten (15) bzw. Anwendungsgebieten unterteilt werden.

16. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 7, dadurch gekennzeichnet, dass die Datensatzeinträge des Alternativen-Lexikons nach Autoren (16) des zugrundeliegenden gesprochenen oder korrigierten Textes unterteilt werden.

20 17. Alternativen-Lexikon-Erstellungsverfahren nach Anspruch 7, dadurch gekennzeichnet, dass das Alternativen-Lexikon (23) während der Korrektur von erkannten Texten online adaptiert wird.

Zusammenfassung:Spracherkennungs- und Korrektursystem

- 5 Bei einem Spracherkennungs- und Korrektursystem, das zumindest eine Spracherkennungseinrichtung (1), der ein gesprochener Text (GT) zuführbar und in einen erkannten Text (ET) transkribierbar ist, sowie eine mit der zumindest einen Spracherkennungseinrichtung (1) über ein Datennetz (2) zur Übertragung des erkannten Textes (ET) und gegebenenfalls des gesprochenen Textes (GT) verbundene
- 10 Korrekturereinrichtung (3) zum Korrigieren des von der zumindest einen Spracherkennungseinrichtung (1) erkannten Textes (ET) umfasst, weist die Korrekturereinrichtung (3) ein Alternativen-Lexikon (23) auf, das Wortteile, Wörter bzw. Wortfolgen enthält, die von der Korrekturereinrichtung (3) als Alternativen zu einzelnen Wortteilen, Wörtern bzw. Wortfolgen des erkannten Textes anzeigbar (22) sind.
- 15 (Figur 1)

1/1

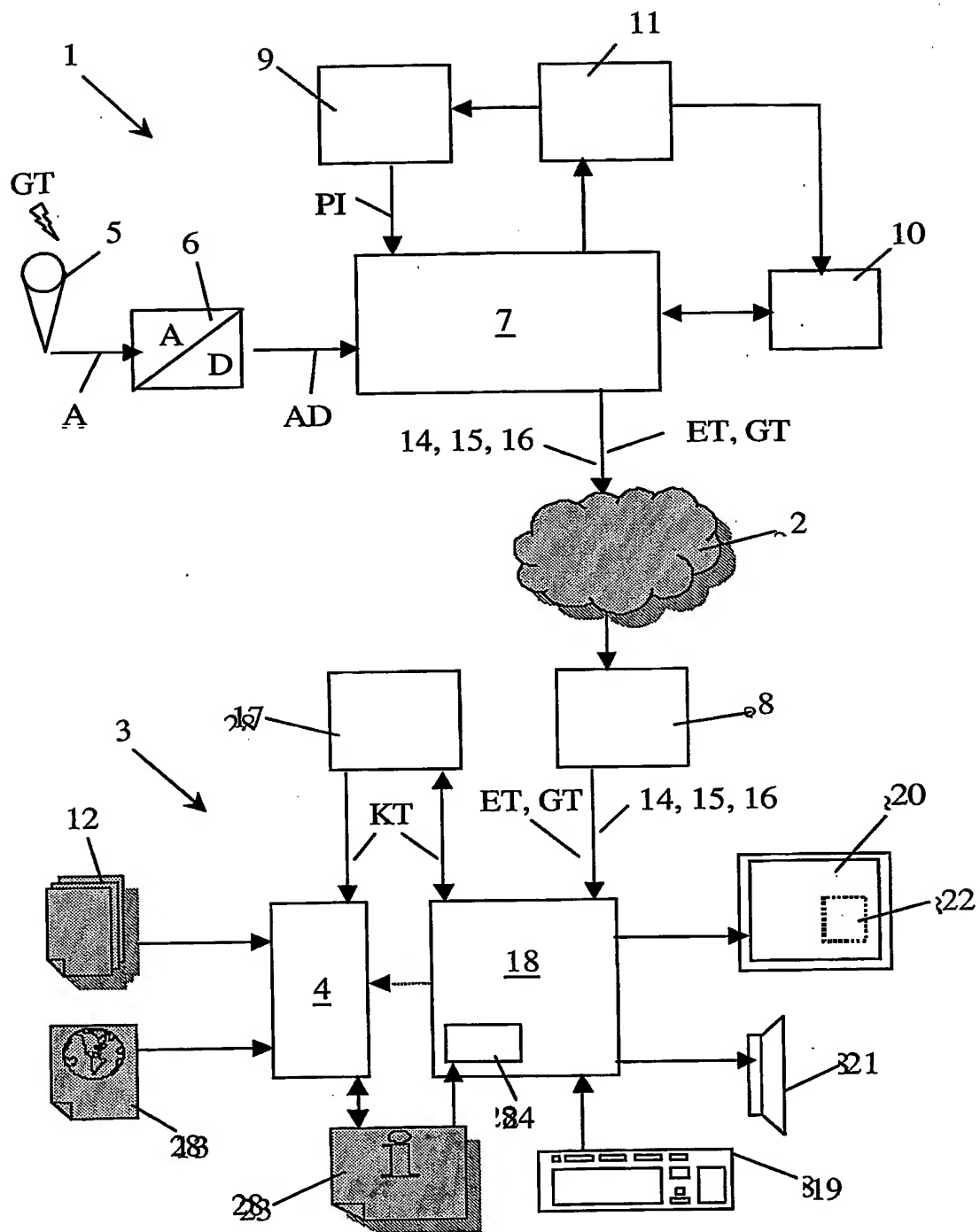


Fig.1

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.